THE INTERRELATIONSHIP OF STRATEGIC OBJECTIVES[1]

Daniel Ellsberg

---

[1] This paper is considerably condensed from P-2183, "The Interrelationship of Strategic Objectives," The RAND Corporation, 1700 Main St., Santa Monica, California; P-2183 develops the present approach further and applies it to a number of proposals and propositions in current debate.

(footnotes in back, double-spaced, serially numbered.

(10)

Many of the significant implications (for US military objectives) of specific military choices--such as the introduction of a new weapons systems, a change in basing or deployment, new operational procedures or protective measures--depend upon their impact on ~~a limited~~ the limited set of US and Soviet "decision elements" shown in the following diagram:

$$
\begin{array}{llcc}
 & & \multicolumn{2}{c}{\text{Soviet Union}} \\
 & & \text{Wait} & \text{Strike} \ (p) \\
\text{US} & \text{Wait} & u_{11}, v_{11} & u_{12}, v_{21} \\
 & \text{Strike} (q) & u_{21}, v_{12} & \text{------}
\end{array}
$$

We shall not try to define a "game" corresponding to this schema, though the format might suggest that interpretation; rather, it depicts some major, interrelated, elements in two concurrent US and Soviet "decision problems": whether or not to launch an all-out nuclear attack upon the opponent "Strike" denotes such an attack; the "Wait" strategy may be interpreted either as a "representative" or as a "best" alternative to Strike. The inclusion of a US Strike strategy ~~need~~ (does) not imply active consideration of such ~~a strategy~~ an alternative at any given moment; as a possibility, it is relevant particularly to Soviet calculations, for reasons indicated below.

The u's and v's are, respectively, US and Soviet "von Neumann-Morgenstern utilities"[2] for certain (highly aggregated) consequences of their ~~actions~~ actions; "p" and "q" are, respectively, the US and Soviet subjective probabilities (expectations, estimates of likelihood) (that a choice of Wait will encounter) ~~for~~ an opponent's choice of Strike during a certain time period. Though all of these are "subjective" variables, they clearly depend upon estimates of objective outcomes under specified circumstances, based upon some form of explicit or tacit "systems analysis." For purposes of this discussion, the v's may be regarded as US estimates of Soviet utilities, estimates which are not held with perfect confidence.

---

[2] I.e., they indicate not merely order of preference among these outcomes but the decision-maker's preferences among ~~risky~~ "gambles": strategies which offer a set of possible outcomes with given subjective probabilities. It is assumed here that ~~number~~ utility numbers can ~~actually~~ be assigned to outcomes so that the decision-maker's actual choices among strategies can be described as maximizing the "mathematical expectation of utility," ~~on the basis of these~~ the average of these utilities weighted by their respective subjective probabilities.

U_{21} and v_{21} are thus, respectively, US and Soviet utility payoffs for ~~"first"first~~ "strike-first" outcomes: the consequences of a surprise nuclear attack upon the opponent. U_{12} and v_{12} are "strike-second" payoffs, reflecting the consequences of being struck first by the opponent. U_{11} and v_{11} are "no all-out war" payoffs, corresponding to situations in which neither opponent chooses Strike.[3]

The precise effects of a change in military "posture," hardware, policy or plans upon these eight variables (including p and q) are, of course, hard to determine, uncertain, and subject to controversy; nevertheless, rough estimates are often made, and these are, in fact, the basis for most policy recommendations as to choices among military alternatives. Indeed, many such recommendations reflect estimates of effects only upon some subset (e.g., one) of these eight factors. ~~While this~~ The above schema has the advantage of directing attention at least to these eight, gross ~~effects~~ consequences of a military change. A typical, ~~ja~~ major military innovation will affect all of these variables, and in what may be opposing directions for a given or for different military objectives. Within a given set of strategic policy alternatives, "conflicts" may be inescapable; an improvement in terms of one dimension of choice (one military sub-goal) may be unavoidably associated with losses with respect to another. Analyses which ignore several of these dimensions are thus likely to be inadequate. Isolated sub-optimizing processes which overlook ~~related sub-goals~~ "conflicts" and "spill-over effects" among related sub-goals may end by lowering overall military security rather than raising it.

¶ While the above, highly simplified and abstract schema can by no means be regarded as an adequate model for the comparison of US military alternatives, it may represent a minimum framework which is an advance over that implicit in much current discussion.

---

[3] The utility subscripts have been chosen to show corresponding elements in the concurrent, related but separate US and Soviet decision problems; if a "game" formulation were being followed, the subscripts for the v's would be transposed.

Ronews ff

Assuming that it is possible to estimate the effects of a military innovation (e.g., an airborne alert, the introduction of IRBM's in Europe, a fall-out shelter program) upon the factors in this schema,[4] the question arises: what effects, or complexes of effects, are "good"? For practical purposes the overall goal of enhancing military security--reducing the likelihood of major losses from the threat or use of enemy military force--must be broken down into military sub-goals, a list of specific, ~~subordinate xxxxxxx,~~ (strategic) objectives. Some of these correspond directly to elements in our schema. Thus, it is a major US objective to ~~xxxxxxxxx~~ lower p: roughly, to "improve the reliability of deterrence." Likewise, ~~xx~~ there is the goal of raising u_{12}: improving the strike-second outcome ~~xxxxxxx~~ if deterrence should fail. Possible conflicts between these two sub-goals are well known: The best outcome if the enemy struck first might be achieved by surrender; ~~but~~ however, by guaranteeing retaliation (lowering v_{21}) it may be possible to lower p greatly, more than compensating for the ~~low~~ lower u_{12} which is associated with the low v_{21}. But what is the effect upon p of improving u_{12}, not by ~~xx~~ planning surrender, but by planning counterforce tactics, or introducing civil defense? To answer this sort of question we must look at the impact not only upon u_{12} but upon all the elements in this framework, ~~xxx~~ for p depends upon the whole configuration of factors in rather a complex way.

To the extent that a Soviet Strike represents a deliberate decision~~, based upon a calculation of payoffs and likelihoods,~~ it must reflect the fact that in Soviet calculations of payoffs and likelihoods, ~~Wait appeared preferable to Strike~~ ~~U(Wait) < U(Strike). The goal of the US "deterrence" policy is to ensure that at all times Wait appears preferable to Strike.~~

---

[4] However these estimates are derived, and whether or not they are "reliable," the schema can be helpful in deriving their policy implications, in order to test the "consistency" of given policy ~~xxxxxxxxxxxxxx~~ recommendations with corresponding estimates and objectives.

at some moment Strike appeared preferable to its best alternative. The goal of the US "deterrence" policy is to ensure that this never arises: that at all times Strike appears inferior in Soviet calculations to some alternative ("Wait"). In our schema this condition appears equivalently as:

(1) V(Wait) $>$ V(Strike), where V is the Soviet utility function; or

(2) $(1-q)v_{11} + q \cdot v_{12} - v_{21} > 0$; or

(3) $(v_{11}-v_{21}) - q(v_{11}-v_{12}) > 0$.

Even though US estimates may indicate ~~that this condition holds~~ at a given moment, p may not be 0; some US uncertainty ($p > 0$) may remain, reflecting:

a) ~~the unreliability of US estimates of Soviet calculations at any one moment~~ the possibility that US estimates are critically mistaken;

(b) the possibility that factors affecting Soviet calculations may *change* critically ~~within the~~ relevant period;

(c) the possibility that Soviet behavior may be ~~[illegible]~~ non-calculated, impulsive or erratic, imperfectly coordinated, or subject to "unauthorized actions" by subordinates.

Each of these ~~possibilities~~ likelihoods is likely to be ~~[illegible]~~ smaller, the *larger* the interval, V(Wait)- V(Strike)[5] ~~the~~ Other things being equal, ~~then the larger this interval[4] --i.e.,~~ the "worse" Strike appears relative to its best alternative, ~~--~~ then, the more likely that the Soviets *are* "deterred," the more likely that they will *stay* deterred as payoffs undergo exogenous shifts, and the more care that Soviet decision-makers will take to avoid accidents, false alarms, hasty decisions, unauthorized ~~actions or uncoordina~~ actions, or uncoordinated, unmonitored policies. The size of this interval, then, provides a sub-criterion among military choices on the path towards lower p.

---

[5] A unit interval having been established by the arbitrary assignment, say, of values 0 and 100 to two specified outcomes.

p. It is, in effect, an index of the sensitivity of the Soviet decision to "counter-deterrent" shifts in payoffs (if q is given): such shifts as (a) a drop in the "no all-out war" outcome v_{11} (due to Soviet losses or expectation of losses in a limited war, shifts in prestige, influence or alliances, "cold war" failures, domestic set-backs or uprisings, political rivalries with third parties);

(b) a drop in the Soviet "strike second" outcome v_{12} (due to increased US force size or ability to exploit weaknesses in Soviet warning systems ~~, defenses, counter~~ or defenses, or prospect of US "annihilation tactics" in US first strike); a ~~drop in~~ rise in the Soviet "strike first" outcome v_{21} ~~(an improvement in Soviet~~ (a reduction in US "strike second" or retaliatory capability, due to changes either in US or in Soviet posture, procedures, tactics). The larger the interval, V(Wait)-V(Strike), the larger (in utility terms) ~~such~~ the payoff "disturbances" ~~would~~ required to make Strike appear preferable to Wait. This might be regarded as one index of the "~~stability~~ reliability of deterrence."

Another important index of this ~~stability~~ reliability is the sensitivity of the Soviet decision to shifts in q, the Soviet expectation of a US first strike. ~~The relevance of q follows from the fact that under typical current and near-future conditions of technology and posture, payoffs tend to have the following structure:[1]~~

~~$$v_{11} \quad v_{21} \quad v_{12}.$$~~

~~This is not an irrevocable state of affairs. It is possible to conceive of a technological and strategic balance in which v_{21} v_{12} (so that a Soviet Strike would not a~~

~~[1] Many of the implications of this relationship between the "strike first" and "strike second" outcomes--the second inequality above--are exposed in Albert Wohlstetter's brilliant and authoritative article, "The Delicate Balance of Terror," Foreign Affairs, Vol. 37, No. 2, January, 1959, 211-234. In part, this present approach is an attempt to formalize some of the propositions in Wohlstetter's argument.~~

To understand why ~~p~~ q is relevant at all to the Soviet choice, let us recall the earlier condition of deterrence:

$$V(\text{Wait})-V(\text{Strike}) = (v_{11}-v_{21}) - q(v_{11}-v_{12}) > 0.$$

Since ~~xxx~~ typically $v_{11} > v_{12}$, it follows that a *necessary* condition for deterrence is:

$$(v_{11}-v_{21}) > 0, \text{ or } v_{11} > v_{21}.$$

It cannot be taken for granted that this condition will hold; it does not follow automatically from the existence on both sides of nuclear weapons.[6] But in any case, this condition is not *sufficient*.

(Perhaps) ~~It is one of~~ the most significant aspect of the current strategic balance is that, under typical conditions of technology and posture:[7]

$$v_{21} > v_{12}.$$

no ital. It follows that deterrence can fail $[\![(v_{11}-v_{21})-q(v_{11}-v_{12})<0]\!]$ even though $(v_{11}-v_{21})$ is positive and ~~very~~ large: if q, the Soviet expectation of a US Strike, is sufficiently great.

An important question is: *How high would q have to be to make Strike appear preferable to the Soviets?* A threshhold value $\tilde{q}$, below which the Soviets would be deterred and above which they would prefer Strike, is given by:

$$(v_{11}-v_{21})-\tilde{q}(v_{11}-v_{12}) = 0, \text{ or } \tilde{q} = \frac{v_{11}-v_{21}}{v_{11}-v_{12}}.$$

---

[6] As Albert Wohlstetter has pointed out, US retaliatory power *could* be so vulnerable to a Soviet "no warning" attack as to promise less destruction than the Russians have suffered historically, ~~an~~ whereas the "no war" outcome could, under abnormal conditions, appear very bad indeed. ("The Delicate Balance of Terror," Foreign Affairs, Vol. 37, No. 2, January, 1959, ~~xxx~~ p. 222).

[7] Many of the implications of this relationship between the "strike first" and "strike second" outcomes are exposed in Wohlstetter's brilliant and authoritative article, *op. cit.* In part, the present approach is an attempt to formalize some of the propositions in Wohlstetter's discussion.

We will refer to $\tilde{q}$, that probability of a US Strike which would, with given Soviet payoffs, make the Soviets indifferent between Strike and Wait, as the "critical risk" of a US Strike. This threshhold expectation, defined as a function of Soviet payoffs, seems a highly significant property of the payoff structure. Among the most important consequences of military choices is their impact upon this parameter, which serves as an index of the sensitivity of the Soviet decision to their expectation of being struck.

Extreme vulnerability of the US retaliatory force can imply a low Soviet critical risk. It leads to an extreme advantage of the "strike first" over the "strike second" outcome, with $(v_{11}-v_{12})$ much greater than $(v_{11}-v_{21})$. With the resulting low $\tilde{q}$, the Soviets would find Strike preferable if they attached even moderate likelihood to a future US Strike. This is clearly an undesirable state of affairs; a Soviet Strike could appear a rational response even to ~~a~~ highly ambiguous indications of a US attack, of the sort generated periodically by any warning system. Under the general objective of improving the "reliability ~~"stability~~ of deterrence" it seems desirable to reduce the sensitivity of the Soviet decision to fluctuations in $\underline{q}$; thus, it becomes a subgoal to increase the critical risk q.

~~By a combination of lowering the Soviet "strike first" outcome v_{21} (reducing the vulnerability of the US retaliatory force, by such measures as additions as Polaris submarines, airborne alert, hardened, dispersed or mobile missiles) and, conceivably, raising the "strike second" outcome v_{12} (by preparations to fight a strictly counter-military campaign~~

The principal method of achieving high $\tilde{q}$--implying that the Soviets will not prefer Strike to Wait unless they are very sure of a US Strike--is to reduce the vulnerability of the US retaliatory force by measures which do not improve markedly the US "strike first" capability: e.g., the replacing of highly vulnerable weapons by Polaris submarines, airborne alert, hardened or land mobile missiles. As $\underline{v}_{21}$ is lowered relative to $\underline{v}_{12}$, a situation is approached in which the Soviets would prefer Wait even if they were certain that the US would attack ($\tilde{q} = 1$, corresponding to $\underline{v}_{21} = \underline{v}_{12}$).[8]

A/~~xxxxx~~ (further) sub-goal, towards improving the ~~stability~~ reliability of deterrence and lowering $\underline{p}$, is to lower $\underline{q}$, the Soviet expectation of a US Strike. Most military choices operate directly upon payoffs, US and Soviet, with indirect effects on expectations. Changes in US payoffs will influence $\underline{q}$ by affecting the Soviet image of the US rational incentives to Strike. Just as $\underline{p}$ corresponds to the US estimate of the reliability of US deterrence, $\underline{q}$ is essentially the Soviet estimate of the reliabilty of Soviet deterrence. To lower $\underline{q}$, then, is to ~~increase the Soviet estimate of the reliability of Soviet deterrence~~ increase, in Soviet eyes, indices of the reliability of Soviet deterrence which are analogous to indices of US deterrence: to increase U(Wait)-U(Strike) in US calculations; to increase $\tilde{p} = \frac{u_{11}-u_{21}}{u_{11}-u_{12}}$, the US "critical risk"; to lower $\underline{p}$.[9] This adds ~~three~~ two new criteria of choice (lowering $\underline{p}$ being already included) to our list of military sub-goals.

Having presented some apparatus of analysis at this length, there is little space in which to apply it here. Let us consider one example, by now rather familiar. Suppose that, as is frequently done, lowering the Soviet "Strike first" outcome $\underline{v}_{21}$ were taken as the only significant sub-goal under the objective of improving the reliability of deterrence; and suppose it were proposed to achieve this by emplacing "soft," fixed, slow-reaction IRBM's in Europe. Like any increase in our inventory of offensive weapons, this move would tend to decrease $\underline{v}_{21}$. But only a little; ~~such weapons~~ fixed IRBM's are subject to no-warning attack by large numbers of accurate, high-yield Soviet medium-range missiles and bombers, and their existence would have a small or negligible effect on the expected outcome of a well-planned Soviet Strike. Even so, if other effects were ignored, as they often are, the move would seem desirable on the basis of this one criterion.

---

[8] Conceivably, this result might be nailed home by making $\underline{v}_{21}$ appear worse than v_{12}; suppose that the Soviets were led to fear US "annihilation tactics" with a large ~~xx~~ retaliatory force in case of a Soviet first strike, but were also made aware that the US was preparing for a strictly countermilitary campaign, avoiding cities and aiming at quick cessation, if war should arise under any other circumstances. See Herman Kahn, On Thermonuclear War (Princeton, 1960), pp.162-189.

[9] An interdependence between $\underline{p}$ and $\underline{q}$ emerges here; it has been ably explored under the heading, "The Reciprocal Fear of Surprise Attack," by Thomas Schelling, in The Strategy of Conflict (Harvard, 1960), pp. 207-229. I would suggest, without developing the point here, that this interaction is most significant when $\tilde{p}$ and $\tilde{q}$, the US and Soviet critical risks, are both low.

However, if we ask the impact of this move upon the other factors in our schema, conflicts with other criteria are likely to emerge. The most marked effects of the innovation would probably be:
a) a sharp decrease in $\underline{v}_{12}$, the Soviet "second strike" outcome;
b) a sharp increase in $\underline{u}_{21}$, the US "strike first" outcome. Neither of these effects, at first glance, might seem undesirable in themselves, ~~thus counting~~ to count as "costs." Yet the drop in $\underline{v}_{12}$ relative to $\underline{v}_{21}$ would imply a lower Soviet critical risk $\tilde{q}$; it would take less assurance than before of a US Strike to make a Soviet Strike seem preferable. And meanwhile, the actual Soviet expectation $\underline{q}$ might be increased; for the rise in $\underline{u}_{21}$ relative to $\underline{u}_{12}$ (which would change negligibly) would mean lower US critical risk $\tilde{p}$, and, for given $\underline{p}$, a reduced interval U(Wait)-U(Strike), so that Soviet deterrence would appear less reliable than before.[10]

---

[10] Herman Kahn has emphasized that such an improvement in $\underline{u}_{21}$ may significantly improve $\underline{u}_{11}$, by deterring the Soviets from such acts short of all-out Strike as might "provoke" a US first strike, if the US first strike outcome were sufficiently high. The objections, which I share, to such a policy of "Type II Deterrence" are too lengthy to discuss here. See Kahn, op. cit., pp. 138-144 and passim. At any rate, note that this sub-goal, if accepted, would in this case conflict with the various criteria of the deterrence of a Soviet Strike (Kahn's "Type I Deterrence").

Thus, the several criteria we have considered for the reliability of US deterrence would indicate that this move might be associated with *higher* p than before.[11] Furthermore, Soviet recognition of this effect could lead, via Schelling's "reciprocal fear of surprise attack," to higher q and a further upward pressure on p.

These results are to be contrasted to those of the measures mentioned earlier for reducing the vulnerability of the retaliatory force (~~without increasing~~ raising q by reducing v_{21} relative to v_{12}); moreover, a complex of such measures may be designed to raise u_{12} much more sharply than u_{21}, thus providing "second strike insurance" against the failure of deterrence while at the same time raising $\tilde{p}$, increasing the reliability of Soviet deterrence and lowering Soviet fears of attack.

¶ ~~Likewise,~~ Other specific arms control, civil defense and active defense measures may be examined in terms of our schema; their implications for the various criteria will depend upon their differential effects upon all of the factors discussed. The discovery, in a particular case, that the implications in terms of several of the criteria (sub-goals) *conflict* is not a failure of the approach; on the contrary, it is a signal of the need for closer analysis in that case, for the weighing of criteria, or for the invention of new alternatives which avoid or alleviate the conflict.

It is clear that this simple framework cannot capture all the complexities of strategic choices. It is in no sense a machine for providing answers; at most, it is a machine for asking useful questions, and for preliminary testing of alleged answers. As such, it can be helpful; simple as it is, it is far more flexible and complex than single-variable models implicit in much "literary" discussion. Unfortunately, there ~~is~~ has been an historical tendency on the part of policy-makers to reject the aid of abstract frameworks of the present sort on the grounds that they are "too simplistic," and then to make practical decisions on the basis of much cruder, implicit models.

---

[11] This argument follows Wohlstetter, *op. cit.*, pp. 222-230, particularly p. 229.